UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/715,552	11/17/2000	Bhaskar Ghosh	50277-0406	6519

29989   7590   11/26/2003

HICKMAN PALERMO TRUONG & BECKER, LLP
1600 WILLOW STREET
SAN JOSE, CA 95125

EXAMINER
NGUYEN, ANH T

ART UNIT	PAPER NUMBER
2127	

DATE MAILED: 11/26/2003

3

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *17 November 2000*.

2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1-26* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-26* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☒ The specification is objected to by the Examiner.

10) ☒ The drawing(s) filed on *17 November 2000* is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a) ☐ All b) ☐ Some * c) ☐ None of:

      1. ☐ Certified copies of the priority documents have been received.

      2. ☐ Certified copies of the priority documents have been received in Application No. _____.

      3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

13) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

    a) ☐ The translation of the foreign language provisional application has been received.

14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .

4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: .

DETAILED ACTION

1. This office action is responsive to application 09/715,552, filed 11/17/2000.

2. Claims 1-26 are presented for examination.

Information Disclosure Statement

3. Applicant is respectfully reminded of the duty to fully disclose information under 37 CFR 1.56 all pertinent information and material pertaining to the patentability of applicant's claimed invention.

Drawings

4. Figure 1 should be designated by a legend such as --Prior Art-- because only that which is old is illustrated. See MPEP § 608.02(g). A proposed drawing correction or corrected drawings are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

Specification

5. The disclosure is objected to because of the following informalities:

   Page 1, line 15, "keys" should recite --key--.

   Appropriate correction is required.

6. The disclosure is objected to because of the following informalities:

   Page 5, line 14, "According to aspect" should recite --According to *one* aspect--.

   Appropriate correction is required.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

7.   Claims 1-4, 13-17, and 26 are rejected under 35 U.S.C. 102(b) as being

anticipated by Gal et al., USPN 5,729,732 (hereinafter Gal).

Regarding Claim 1:

Gal discloses a method for distributing data items from a particular set of data into a

plurality of buckets based on distribution keys associated with said data items (Gal, see

Abstract, "A method is described for operating a computer to sort a set of data records

each having an associated key for governing the sort process"), the method comprising

the steps of:

randomly selecting data items from said particular set of data to produce a

sampled set of data items(Gal, col.3, lines 51-52, "The random sampling can be achieved,

for example, by taking a predetermined set of n indices");

determining a plurality of ranges based on the distribution keys associated

with the sampled set of data items (Gal, see Abstract, "determining a range for the key

values by sampling the key values");

assigning said plurality of ranges to said plurality of buckets (Gal, see Abstract,

"defining a plurality of buckets, each bucket corresponding to a respective one of a

plurality M of subintervals in the range"); and

distributing each data item in said particular set of data to the bucket that has

been assigned the range into which falls the distribution key of the data item (Gal, see

Abstract, "distributing the keys among the buckets by determining directly from each key

value the index of the subinterval into which the key value falls").

Regarding Claim 2:

Gal discloses randomly selecting data items from each subset of a plurality of subsets of

said particular set of data (Gal, col.3, lines 40-41 and 46, "The file to be sorted comprises

N records each of which has an associated key" and "A random sample of the keys is

taken from the file y.sub.1, y.sub.2,. .. y.sub.n").

Regarding Claim 3:

Gal discloses randomly selecting data items from each partition of a partitioned table

(Gal, col.3, lines 51-54, "The random sampling can be achieved, for example, by taking a

predetermined set of n indices, for example from a pseudo-random table, and picking the

corresponding elements").

Regarding Claim 4:

Gal discloses randomly selecting data items from subsets of data, stored in buffers in

volatile memory (i.e. RAM), that represent results of one or more previously performed

operations (Gal, see FIG.1, col.3, lines 12-13, "the data processing system which may be

utilized for implementing the method and system of the present invention includes a

processor 10, a random access memory (RAM) 12, a read only memory (ROM) 14, at

least one non-volatile storage device 15, a computer display monitor 16 and a keyboard

18").

Regarding Claim 13:

Gal discloses determining ranges that contain an approximately equal amount of

distribution keys associated with said sampled set of data items (Gal, col.2. lines 7-19,

"the index of the subinterval into which each key falls is determined directly from the key

value. This means that the distribution of each key into the respective bucket can be

performed in a time, which does not depend on the number of buckets used in the

distribution. The subintervals are equal").

8.   Claims 14- 17 and 26 are claims to a computer readable medium carrying

instructions, which performs the steps of the method of claims 1-4 and 13. Computer

readable medium include CDs, floppy disks, hard drives, memory, etc. Gal teaches a

computer implemented process, thus it is inherent that the program accomplishing the

procedures must be carried or stored on a computer readable medium to enable the

computer to function in the manner taught by Gal. Therefore, claims 14-17 and 26 are

rejected for the reasons set forth above and under the same rationale as claims 1-4 and 13.

Regarding Claim 14:

Gals discloses:

distributing data items from a particular set of data into a plurality of buckets

based on distribution keys associated with said data items (Gal, see Abstract, "A method

is described for operating a computer to sort a set of data records each having an

associated key for governing the sort process"),

randomly selecting data items from said particular set of data to produce a

sampled set of data items (Gal, col.3, lines 51-52, "The random sampling can be

achieved, for example, by taking a predetermined set of n indices");

determining a plurality of ranges based on the distribution keys associated

with the sampled set of data items (Gal, see Abstract, "determining a range for the key

values by sampling the key values");

assigning said plurality of ranges to said plurality of buckets (Gal, see Abstract,

"defining a plurality of buckets, each bucket corresponding to a respective one of a

plurality M of subintervals in the range"); and

distributing each data item in said particular set of data to the bucket that has

been assigned the range into which falls the distribution key of the data item (Gal, see

Abstract, "distributing the keys among the buckets by determining directly from each key

value the index of the subinterval into which the key value falls").

Regarding Claim 15:

Gal discloses randomly selecting data items from each subset of a plurality of subsets of

said particular set of data (Gal, col.3, lines 40-41 and 46, "The file to be sorted comprises

N records each of which has an associated key" and "A random sample of the keys is

taken from the file y.sub.1, y.sub.2,. . . y.sub.n").

Regarding Claim 16:

Gal discloses randomly selecting data items from each partition of a partitioned table

(Gal, col.3, lines 51-54, "The random sampling can be achieved, for example, by taking a

predetermined set of n indices, for example from a pseudo-random table, and picking the

corresponding elements").

Regarding Claim 17:

Gal discloses randomly selecting data items from subsets of data, stored in buffers in volatile memory, that represent results of one or more previously performed operations (Gal, see FIG.1, col.3, lines 12-13, "the data processing system which may be utilized for implementing the method and system of the present invention includes a processor 10, a random access memory (RAM) 12, a read only memory (ROM) 14, at least one non-volatile storage device 15, a computer display monitor 16 and a keyboard 18").

Regarding Claim 26:

Gal discloses determining ranges that contain an approximately equal amount of distribution keys associated with said sampled set of data items (Gal, col.2. lines 7-19, "the index of the subinterval into which each key falls is determined directly from the key value. This means that the distribution of each key into the respective bucket can be performed in a time, which does not depend on the number of buckets used in the distribution. The subintervals are equal").

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

9.   Claims 5, 7, 12, 18, 20, and 25 rejected under 35 U.S.C. 103(a) as being

unpatentable over Gal et al., USPN 5,729,732 (hereinafter Gal) in view of Ogi (USPN

5,854,938).

Regarding Claim 5:

Gal discloses a method for evenly distributing data items to corresponding

buckets (Gal, see Abstract, "A method is described for operating a computer to sort a set

of data records each having an associated key for governing the sort process, the method

comprising determining a range for the key values by sampling the key values; defining a

plurality of buckets, each bucket corresponding to a respective one of a plurality M of

subintervals in the range").

However, Gal does not particularly disclose processing the buckets with plural

processors concurrently operating in parallel to execute a task.

Ogi discloses assigning the plurality of buckets to a plurality of processes (Ogi,

see FIG.3); and causing each process of said plurality of processes to perform, in parallel

with the other processes of said plurality of processes, an operation on the data items

contained in any buckets assigned to the process (Ogi, see FIG.3).

Gal and Ogi are analogous art because they are from the same field of endeavor in

parallel operations and data distribution. It would have been obvious for one of ordinary

skill in the art at the time the invention was made to modify the method for sorting a set

of data in a computer system using buckets having associated key values within a range

of Gal such that it utilizes the parallel computer system having plural processors

concurrently operating in parallel to execute tasks as disclosed by Ogi.

One of ordinary skill in the art would be motivated to do so because it increases

data management efficiency by optimally storing data in a plurality of database storage

areas and decreasing the time it takes for a computer system to perform a plurality of

tasks. This improvement over the prior art would result in efficient load balancing

thereby optimizing performance in database systems.

Regarding Claim 7:

Ogi discloses the particular set of data is durably stored on a plurality of durable storage

units (ogi, see FIG.3, element 24); and randomly selecting durable storage units from said

plurality of durable storage units and using the data items stored on said randomly

selected durable storage units as the sampled set of data items (Ogi, see FIG.3, col.14,

lines 41-43, "Inside each of the bucket groups, plural buckets (of n sorts here) are stored

at random regardless of the order of generating the bucket in a level of tuples").

Regarding Claim 12:

Ogi discloses wherein said operation is specified in a database command,

the method further comprising receiving with said database command data that

indicates how much of said particular set of data to randomly select to produce said

sampled set of data items (Ogi, col.3, lines 30-37, "Meanwhile, a work of dividing a tuple

group into plural groups (buckets) using a grouping function (classification using a value

of hashing or a specific column, for example) is extremely general in an RDB process.

For instance, "Group By" phrase in the SQL statement is to clearly demand for a group

dividing process, or hash join which is a typical system of join operations with the above

structure features group sorting of tuples with a hash function").

10.   Claims 18, 20 and 25 are claims to a computer readable medium carrying

instructions, which performs the steps of the method of claims 5, 7 and 12. Computer

readable medium include CDs, floppy disks, hard drives, memory, etc. Gal teaches a

computer implemented process, thus it is inherent that the program accomplishing the

procedures must be carried or stored on a computer readable medium to enable the

computer to function in the manner taught by Gal. Therefore, claims 18, 20 and 25 are

rejected for the reasons set forth above and under the same rationale as claims 5, 7 and

12.

Regarding Claim 18:

Ogi discloses assigning the plurality of buckets to a plurality of processes (Ogi, see

FIG.3); and causing each process of said plurality of processes to perform, in parallel

with the other processes of said plurality of processes, an operation on the data items

contained in any buckets assigned to the process (Ogi, see FIG.3).

Regarding Claim 20:

 Ogi discloses the particular set of data is durably stored on a plurality of durable storage

units (ogi, see FIG.3, element 24); and randomly selecting durable storage units from said

plurality of durable storage units and using the data items stored on said randomly

selected durable storage units as the sampled set of data items (Ogi, see FIG.3, col.14,

lines 41-43, "Inside each of the bucket groups, plural buckets (of n sorts here) are stored

at random regardless of the order of generating the bucket in a level of tuples").

Regarding Claim 25:

Ogi discloses wherein said operation is specified in a database command,

the method further comprising receiving with said database command data that

indicates how much of said particular set of data to randomly select to produce said

sampled set of data items (Ogi, col.3, lines 30-37, "Meanwhile, a work of dividing a tuple

group into plural groups (buckets) using a grouping function (classification using a value

of hashing or a specific column, for example) is extremely general in an RDB process.

For instance, "Group By" phrase in the SQL statement is to clearly demand for a group

dividing process, or hash join which is a typical system of join operations with the above

structure features group sorting of tuples with a hash function").

11.   Claims 6, and 19 rejected under 35 U.S.C. 103(a) as being unpatentable over the

prior art as applied to claim 1 above and further in view of Marks USPN 5,748,844.

Regarding Claim 6:

Gal discloses a method for randomly selecting data items from each subset of a

plurality of subsets of said particular set of data (Gal, col.3, lines 40-41 and 46, "The file

to be sorted comprises N records each of which has an associated key" and "A random

sample of the keys is taken from the file y.sub.1, y.sub.2,. .. y.sub.n").

However, Gal does not particularly disclose selecting a distinct random seed

for each subset of the plurality of subsets of said particular set of data.

Marks discloses selecting a distinct random seed for each subset of the plurality of

subsets of said particular set of data (Marks, FIG.3, element 32, col.3, lines 44-46 "the

seed-growth heuristic initially assigns a small number of randomly chosen nodes to each

part of the partition; these are the seed nodes").

Gal and Marks are analogous art because they are from the same field of endeavor

in partitioning that can be applied to database design and parallel processing. Generally,

random number generators have to be "seeded" with an initial seed. The use of random

seed generators from which a "first" random number is derived is well known in the art.

It would have been obvious for one of ordinary skill in the art at the time the

invention was made to modify the method for randomly selecting data items from each

subset of a plurality of subsets such that it utilizes the random seed generator of Marks.

One of ordinary skill in the art would be motivated to do so because it assures the

best random sampling of data for partitioning thereby yielding a superior sampling

operation performed in parallel.

Regarding Claim 19:

Claim 19 is a claim to a computer readable medium carrying instructions, which performs

the steps of the method of claim 6. Computer readable medium include CDs, floppy

disks, hard drives, memory, etc. Gal teaches a computer implemented process, thus it is

inherent that the program accomplishing the procedures must be carried or stored on a

computer readable medium to enable the computer to function in the manner taught by

Gal. Therefore, claim 19 is rejected for the reasons set forth above and under the same

rationale as claim 6.

12.   Claims 8, 10, 21, and 23 rejected under 35 U.S.C. 103(a) as being unpatentable

over the prior art as applied to claim 1 above, and in further view of Couch et al. USPN

6,604,096 (hereinafter Couch).

Regarding Claim 8:

Gal discloses randomly selecting data items from said particular set of data to produce a sampled set of data items(Gal, col.3, lines 51-52, "The random sampling can be achieved, for example, by taking a predetermined set of n indices").

Gal does not particularly disclose selecting a specified *percentage* of data items in said particular set of data.

Couch discloses selecting a specified percentage of data items in said particular set of data (Couch, see FIG.6, element 120, col.9, lines 49-51, "percentage designation control 120 may be use to select the amount of data").

Gal and Couch are analogous art because they are from the same field of endeavor in database management systems. It would have been obvious for one of ordinary skill in the art at the time the invention was made to implement the method for randomly selecting data items from said particular set of data to produce a sampled set of data items such that it incorporates selecting a specified percentage of data items in said particular set of data as disclosed by Marks.

One of ordinary skill in the art would be motivated to do so because it assures the even distribution of data resulting in improved quality of partitioning.

Regarding Claim 10:

Couch discloses the method of Claim 8 further comprising the step of receiving, from a user, data that specifies said percentage (Couch, col.14, lines 1-2, "configured to receive a user selection of a percentage of the data").

13.   Claims 21 and 23 are claims to a computer readable medium carrying

instructions, which performs the steps of the method of claims 8 and 10. Computer

readable medium include CDs, floppy disks, hard drives, memory, etc. Gal teaches a

computer implemented process, thus it is inherent that the program accomplishing the

procedures must be carried or stored on a computer readable medium to enable the

computer to function in the manner taught by Gal. Therefore, claims 21 and 23 are

rejected for the reasons set forth above and under the same rationale as claims 8 and 10.

Regarding Claim 21:

        Gal discloses randomly selecting data items from said particular set of data to

produce a sampled set of data items(Gal, col.3, lines 51-52, "The random sampling can

be achieved, for example, by taking a predetermined set of n indices").

        Gal does not particularly disclose selecting a specified *percentage* of data items in

said particular set of data.

        Couch discloses selecting a specified percentage of data items in said particular

set of data (Couch, see FIG.6, element 120, col.9, lines 49-51, "percentage designation

control 120 may be use to select the amount of data").

        Gal and Couch are analogous art because they are from the same field of endeavor

in database management systems. It would have been obvious for one of ordinary skill in

the art at the time the invention was made to implement the method for randomly

selecting data items from said particular set of data to produce a sampled set of data items

such that it incorporates selecting a specified percentage of data items in said particular

set of data as disclosed by Marks.

One of ordinary skill in the art would be motivated to do so because it assures the

even distribution of data resulting in improved quality of partitioning.

Regarding Claim 23:

Couch discloses the method of Claim 8 further comprising the step of receiving, from a

user, data that specifies said percentage (Couch, col.14, lines 1-2, "configured to receive

a user selection of a percentage of the data").

14.   Claims 9, 11, 22, and 24 are rejected under 35 U.S.C. 103(a) as being

unpatentable over the prior art as applied to claim 7 above, and further in view of Couch

et al. USPN 6,604,096 (hereinafter Couch).

Regarding Claim 9:

Keeping in mind what Gal and Ogi discloses as mentioned above, the reference of

Gal modified by Ogi does not teach selecting a specified percentage of the plurality of

durable storage units that are storing said particular set of data.

Couch discloses selecting a selecting a specified percentage of the plurality of

durable storage units that are storing said particular set of data (Couch, col.5, lines 13-15,

"The operational data may be collected as a single data set, or may be distributed over

different locations including over different storage devices").

It would have been obvious for one of ordinary skill in the art at the time the

invention was made to implement the method for randomly selecting data items from said

particular set of data to produce a sampled set of data items such that it incorporates

selecting a selecting a specified percentage of the plurality of durable storage units that

are storing said particular set of data as disclosed by Marks.

One of ordinary skill in the art would be motivated to do so because it allows

flexible storage of data thereby improving performance of the database management

system.

Regarding Claim 11:

Couch discloses the method of Claim 9 further comprising the step of receiving , from a

user, data that specifies said percentage (Couch, col.14, lines 1-2, "configured to receive

a user selection of a percentage of the data").

15.   Claims 22 and 24 are claims to a computer readable medium carrying

instructions, which performs the steps of the method of claims 9 and 11. Computer

readable medium include CDs, floppy disks, hard drives, memory, etc. Gal teaches a

computer implemented process, thus it is inherent that the program accomplishing the

procedures must be carried or stored on a computer readable medium to enable the

computer to function in the manner taught by Gal. Therefore, claims 22 and 24 are

rejected for the reasons set forth above and under the same rationale as claims 9 and 11.

Regarding Claim 22:

Keeping in mind what Gal and Ogi discloses as mentioned above, the reference of

Gal modified by Ogi does not teach selecting a specified percentage of the plurality of

durable storage units that are storing said particular set of data.

Couch discloses selecting a selecting a specified percentage of the plurality of

durable storage units that are storing said particular set of data (Couch, col.5, lines 13-15,

"The operational data may be collected as a single data set, or may be distributed over

different locations including over different storage devices").

It would have been obvious for one of ordinary skill in the art at the time the

invention was made to implement the method for randomly selecting data items from said

particular set of data to produce a sampled set of data items such that it incorporates

selecting a selecting a specified percentage of the plurality of durable storage units that

are storing said particular set of data as disclosed by Marks.

One of ordinary skill in the art would be motivated to do so because it allows

flexible storage of data thereby improving performance of the database management

system.

Regarding Claim 24:

Couch discloses the method of Claim 9 further comprising the step of receiving, from a

user, data that specifies said percentage (Couch, col.14, lines 1-2, "configured to receive

a user selection of a percentage of the data").

Conclusion

16.   The prior art made of record and not relied upon is considered pertinent to applicant's

disclosure. See Form PTO-892.

17.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Anh T Nguyen whose telephone number is (703) 305-8649.

The examiner can normally be reached on Monday-Friday from 7:00 AM to 4:30 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, William Grant, can be reached on (703) 308-1108. The fax phone number for

the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding

should be directed to the receptionist whose telephone number is (703) 306-5484.

Anh T.Nguyen

Art Unit 2127

November 24, 2003

WILLIAM GRANT
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100